Verklarende statistiek in SPSS

1. Vergelijken van gemiddelden voor 2 groepen

1.1 Ongepaarde t-test

Levene's Test for equality of variances:

- H_0 varianties zijn gelijk
- $p < 0.05$ (5%) => H_0 verwerpen => varianties niet gelijk
- $p > 0.05$ (5%) => geen bewijs dat de varianties gelijk zijn => t-test Welch benadering voor de zekerheid

De constructie van de teststatistiek van een ongepaarde t-test gaat ervan uit dat elke groep normaal verdeeld is. => elke groepsgemiddelde normaal verdeeld.

➔ Verdelingen van de groepen niet normaal verdeeld => het steekproefgemiddelde van elke groep zal bij benadering normaal verdeeld zijn indien de steekproefgrootte N>30

De ongepaarde t-test gaat ervan uit dat de groepen onafhankelijk zijn van elkaar.

Normaliteit van elke groep na te gaan => QQ-plots en Shapiro-Wilk test

QQ-plot: data normaal verdeeld => punten op een rechte lijn

Shapiro-Wilk test:

- H_0 data normaal verdeeld
- $p < 0.05$ (5%) => H_0 verwerpen => niet normaal verdeeld MAAR N>30 toch uitvoeren
- $p > 0.05$ (5%) => H_0 niet verwerpen => normaal verdeeld

<u>Voorbeeld tweezijdige ongepaarde t-test:</u>

Scoren de mannelijke studenten gemiddeld beter of slechter op het vak Wiskunde II (A) in vergelijking met de vrouwelijke studenten?

Wanneer we de gemiddelden voor beide groepen opvragen via **Compare Means → Means** zien we dat het steekproefgemiddelde bij de mannen gelijk is aan 11, 00 en bij de vrouwen 8, 63. We willen echter niet enkel weten of de gemiddelden in de steekproef verschillen, maar ook of we deze verschillen mogen veralgemenen naar de populatie (in dit geval bestaat de populatie uit alle studenten EW/TEW/HIR waaruit de steekproef is genomen). Daarom zullen we een ongepaarde t-test uitvoeren.

Eerst zullen we de normaliteit van de groepen nagaan. Ga hiervoor naar **Analyze → Descriptive Statistics → Explore**. Kies als *Dependent List* Wiskunde II (A) en als *Factor List* de variabele geslacht. Kies bij *Display* de optie *Plots* en bij de knop *Plots* vink je de optie *Normality plots with tests* aan en vink je de andere opties uit of selecteer je *None*. Druk op *Continue* en *OK*.

In de output zien we de p-waarden voor de Shapiro-Wilk test en de QQ-plots. Voor de mannelijke studenten geeft de Shapiro-Wilk test een p-waarde aan van 0, 157. In de bijhorende QQ-plot zien we dat de (weinige) punten ongeveer op een rechte liggen, enkel links zien we een punt dat wat verder

afwijkt van de rechte. Op basis hiervan besluiten we dat de gegevens mij de mannen voldoende normaal verdeeld zijn.

Voor de vrouwelijke studenten is de p-waarde gelijk aan 0, 031. *De afwijking van normaliteit is hier dus wel significant*. Ook in de QQ-plot zien we dat er meerdere punten verder afwijken van de rechte. De voorwaarde van normaal verdeelde gegevens is dus niet voldaan voor de gegevens bij de vrouwen. In principe zouden we beter een niet-parametrische test uitvoeren, maar deze zullen we niet behandelen in SPSS. We zullen dus toch een t-test uitvoeren, maar moeten in het achterhoofd houden dat de voorwaarden eigenlijk niet voldaan zijn.

De t-test zelf kunnen we uitvoeren via **Analyze → Compare Means → Independent Samples T Test.** Kies als Test Variable Wiskunde II (A) en als Grouping Variable de variabele geslacht. Klik vervolgens op de knop Define Groups en geef voor groep 1 de waarde M, voor groep 2 de waarde V en druk op Continue en OK. In de output vinden we bij Group Statistics onder meer de steekproefgemiddeldes terug.

Bij Independent Samples Test vinden we links de output van de Levene's Test for Equality of Variances, die aantoont dat er geen significant verschil is tussen de varianties in beide groepen (p = 0, 721). Zoals reeds gezegd is dit echter geen bewijs dat de varianties wel gelijk zijn en zullen we voor de veiligheid kiezen voor de Welch benadering. De p-waarde van deze test vinden we op de onderste lijn (Equal variances not assumed) en is gelijk aan 0, 073.

We kunnen dus op het 5% significantie-niveau niet besluiten dat de gemiddelde score op het vak Wiskunde II (A) verschilt tussen mannelijke en vrouwelijke studenten. Merk op dat we wel een significant verschil zouden gevonden hebben indien we de t-test met gepoolde steekproefstandaarddeviatie gekozen zouden hebben. De p-waarde daarvan is immers p = 0, 048, zoals je kan aflezen op de bovenste rij (Equal variances assumed).

1.2 Eenzijdig testen

Eenzijdig testen => manueel eenzijdige p-waarde afleiden uit de output van SPSS.

Stel T de teststatistiek, tk de t-verdeling die T volgt als de nulhypothese waar is en t de waarde van de teststatistiek voor deze steekproef, dan is de tweezijdige p-waarde gelijk aan

$p = P(T < -|t| \cup T > |t|) = P(T < -|t|) + P(T > |t|) = 2P(T < -|t|) = 2P(T > |t|)$

Stel dat we groep A en groep B hebben. Het teken van t is dan hetzelfde als het teken van $\bar{x}_A - \bar{x}_B$

Stel dat we willen testen of het gemiddelde van groep A groter is dan het gemiddelde van groep B. De nulhypothese luidt dan **μA = μB** en de alternatieve hypothese luidt **μA > μB** . De eenzijdige p-waarde is dan gelijk aan de kans dat de teststatistiek T minstens even groot is als t gegeven dat de nulhypothese (μA = μB) waar is, m.a.w. P(T ≥ t).

- Indien $\bar{x}_A > \bar{x}_B$, dan is t positief en ondersteunt de steekproef de alternatieve hypothese. De tweezijdige p-waarde is dan p = 2P(T > |t|) = 2P(T > t), m.a.w. is de eenzijdige p-waarde P(T > t) gelijk aan $\frac{p}{2}$.

– Indien $\bar{x}_A < \bar{x}_B$, dan is t negatief en ondersteunt de steekproef de alternatieve hypothese niet. De tweezijdige p-waarde is dan p = 2P (T < −|t|) = 2P (T < t) = 2(1−P(T > t)). M.a.w. is de eenzijdige p-waarde P(T > t) gelijk aan $1 - \frac{p}{2}$.

Stel dat we willen testen of het gemiddelde van groep A kleiner is dan het gemiddelde van groep B. De nulhypothese luidt dan **μA = μB** en de alternatieve hypothese luidt **μA < μB** . De eenzijdige p-waarde is dan gelijk aan de kans dat de teststatistiek T minstens even klein is als t gegeven dat de nulhypothese (μA = μB) waar is, in dit geval P(T < t).

- Indien $\bar{x}_A > \bar{x}_B$, dan is t positief en ondersteunt de steekproef de alternatieve hypothese niet. De tweezijdige p-waarde is dan p = 2P(T > |t|) = 2P(T > t) = 2(1−P(T < t)), m.a.w. is de eenzijdige p-waarde P(T < t) gelijk aan $1 - \frac{p}{2}$.

- Indien $\bar{x}_A < \bar{x}_B$, dan is t negatief en ondersteunt de steekproef de alternatieve hypothese. De tweezijdige p-waarde is dan p = 2P (T < −|t|) = 2P (T < t). M.a.w. is de eenzijdige p-waarde P(T < t) gelijk aan $\frac{p}{2}$.

Samenvattend: indien de steekproef de alternatieve hypothese ondersteunt, dan is de eenzijdige p-waarde de helft van de tweezijdige p-waarde. Als de steekproef de alternatieve hypothese niet ondersteunt, dan is de eenzijdige p-waarde gelijk aan één min de helft van de tweezijdige p-waarde.

Voorbeeld eenzijdige ongepaarde t-test

Beschouw de volgende onderzoeksvraag:

Scoren studenten Handelsingenieur beter op Financiële Analyse van de Onderneming (FAO) dan de andere studenten?

Als nulhypothese hebben we dat de gemiddelde score op FAO van de studenten HIR gelijk is aan de gemiddelde score op FAO van de studenten TEW en EW. De alternatieve hypothese luidt dat de gemiddelde score op FAO van de studenten HIR hoger is dan de gemiddelde score op FAO van de studenten TEW en EW.

Voor deze oefening zullen we gebruik maken van de variabele richting2, die we aangemaakt hebben in de vorige les. We zullen eerst nagaan of de twee groepen normaal verdeeld zijn. Ga naar **Analyze → Descriptive Statistics → Explore**, vul bij *Dependent List* de variabele Financiële Analyse van de Onderneming in en bij *Factor List* richting2. Kies bij *Display* de optie *Plots* en vink bij *Plots* de optie *normality plots with tests* aan en kies bij al de rest None of vink uit. Druk op *Continue* en *OK*.

We vinden voor de normaliteitstest van FAO bij HIR een p-waarde van 0, 156, m.a.w. er is geen bewijs dat de verdeling afwijkt van een normale verdeling. Bij de normaliteitstest van FAO bij (T)EW vinden we een p-waarde van 0, 364. Ook hier is er onvoldoende bewijs om de nulhypothese te verwerpen op het 5% significantieniveau. De Normal QQ-plots bevestigen dit. We gaan er dus van uit dat we de t-test mogen uitvoeren.

Om de t-test zelf uit te voeren gaan we opnieuw naar **Analyze → Compare Means → Independent Samples T-test,** vul bij *Test variable* Financiële Analyse van de Onderneming in en bij *Grouping Variable* richting2 in. Er wordt bovendien gevraagd de groepen te definiëren (Define Groups), voer hier bijvoorbeeld voor *Group 1* HIR en voor *Group 2* (T)EW in.

We kiezen standaard voor de t-test met Welch benadering. In dit geval wordt deze keuze ondersteund door de Levene's Test for Equality of Variances, omdat de nulhypothese van gelijke varianties verworpen kan worden (p=0, 010). De t-test geeft een p-waarde van 0, 953 (tweezijdig).

Om de eenzijdige p-waarde voor deze alternatieve hypothese te berekenen, kijken we eerst naar de steekproefgemiddeldes van de FAO per groep (HIR en (T)EW). Bij de studenten HIR is dit 13, 73 en bij de studenten (T)EW is dit 13, 79. Het steekproefgemiddelde bij de Handelsingenieurs is dus kleiner dan het steekproefgemiddelde bij de (T)EW. Deze steekproefgemiddelden ondersteunen de alternatieve hypothese dus niet, m.a.w. is de eenzijdige p-waarde gelijk aan 1 − 0,953 2 = 0, 5235. *We besluiten dat de gemiddelde score op Wiskunde II (A) van de studenten HIR niet significant groter is dan de gemiddelde score op FAO van de studenten (T)EW op het 5% significantieniveau.*

1.3 Gepaarde t-test

Om een gepaarde t-test uit te mogen voeren, moeten de verschillen van de gepaarde gegevens normaal verdeeld zijn of hebben we voldoende observaties nodig opdat het gemiddelde normaal verdeeld is dankzij de centrale limietstelling.

Om de normaliteit te testen, moeten we eerst een variabele aanmaken die gelijk is aan de verschillen van de twee variabelen.
Via **Transform → Compute Variable** kunnen we dit verschil tussen de gepaarde observaties aanmaken en kunnen we daarna de methode **Analyze → Descriptive statistics → Explore** opnieuw gebruiken (deze keer zonder grouping variable).
Hierna kan je de gepaarde t-test uitvoeren via **Analyze → Compare Means → Paired-Samples T-test**.

Voorbeeld tweezijdige gepaarde t-test

Beschouw de onderzoeksvraag:

Is de verwachte score op Macro-economie verschillend van de verwachte score op Financiële Analyse van de Onderneming?

De **nulhypothese** luidt **μmacro = μfao** en de **alternatieve hypothese** luidt **μmacro ≠ μfao**.

We starten opnieuw met de normaliteit na te gaan. Hiervoor moeten we eerst *de variabele 'verschil.macro.fao' aanmaken*. Ga naar **Transform → Compute Variable**, vul verschil.macro.fao in bij *Target Variable*, macro − fao bij *Numeric Expression* en druk op *OK*. Ga nu naar **Analyze → Descriptive statistics → Explore**, vul bij *Dependent List* verschil.macro.fao in, kies optie *Plots* en selecteer bij *Plots* enkel de optie *Normality plots with tests*. We vinden een p-waarde van 0, 661 voor de normaliteitstest. We besluiten dat de verdeling van de verschillen niet sterk afwijkt van een normale verdeling. We mogen dus een gepaarde t-test uitvoeren.

Ga nu naar **Analyze → Compare Means → Paired-Samples T-test** selecteer Macro-analyse als *Variable1* en Financiële Analyse van de Onderneming als *Variable2*. Druk op *OK*. We zien dat het steekproefgemiddelde 12, 64 bedraagt voor Macro-economie en 13, 76 voor Financiële Analyse van de Onderneming. De gepaarde t-test geeft een p-waarde van 0, 042, we kunnen de nulhypothese dus verwerpen. We besluiten op het 5%-significantieniveau dat de gemiddelde score op Macro-economie niet gelijk is aan de gemiddelde score op Financiële Analyse van de Onderneming.

Voorbeeld eenzijdige gepaarde t-test

Beschouw de volgende onderzoeksvraag:

Is de gemiddelde score op het vak Macro-economie hoger dan de gemiddelde score op Micro-economie?

De **nulhypothese** luidt **μmacro = μmicro** en de **alternatieve hypothese** luidt **μmacro > μmicro**.

We *maken de variabele verschil.macro.micro aan* via **Compute variable**, deze heeft als formule macro−micro. Vervolgens gaan we via **Explore** de normaliteit na van dit verschil. We vinden een p-waarde van 0, 350 voor de normaliteitstest, er is dus geen sterke afwijking van normaliteit, zoals we ook in de QQ-plot kunnen zien. Bij **Paired-Samples T Test** kiezen we Macro-economie als *Variable1*, Micro-economie als *Variable2* en drukken op *OK*. We vinden een tweezijdige p-waarde van 0, 004 voor de gepaarde t-test.

Om de eenzijdige p-waarde te bepalen, kijken we naar de steekproefgemiddelden van Macro-economie en Micro-economie. Bij Macro-economie (12, 64) is dit groter dan bij Micro-economie (11, 52). Dit ondersteunt de alternatieve hypothese, we besluiten dat de eenzijdige p-waarde gelijk is aan 0,004 2 = 0, 002. Er is dus sterk bewijs tegen de nulhypothese, we besluiten dat de gemiddelde score op Macro-economie hoger is dan de gemiddelde score op Micro-economie.

2. Variantieanalyse (ANOVA)

Een **ANOVA test** heeft als *nulhypothese dat alle groepsgemiddelden gelijk zijn* aan elkaar en als *alternatieve hypothese dat minstens 2 groepen verschillen qua gemiddelden*.

Indien we de *nulhypothese verwerpen*, kunnen we een *post hoc analyse uitvoeren* om te onderzoeken welke groepen verschillen van gemiddelde. Hier zullen we de groepsgemiddelden paarsgewijs vergelijken via t-testen.

De ANOVA test gaat uit van:

- normaliteit van de data in elke groep,
- gelijkheid van de varianties van elke groep,
- de onafhankelijkheid van de groepen onderling.

Net zoals bij de ongepaarde t-test blijft het percentage type I fouten quasi gelijk indien één of meer groepen niet normaal verdeeld zijn, maar wel groot genoeg (centrale limietstelling). We zullen hier opnieuw 30 nemen als minimumhoeveelheid. Indien er in een groep sterke afwijkingen van normaliteit zijn en deze groep niet veel meer dan 30 observaties zijn, zullen we eerder terughoudende conclusies trekken bij de ANOVA test wanneer de p-waarde tussen 1% en 5% ligt.

De normaliteitsassumptie kunnen we opnieuw nagaan via **Analyze → Descriptive Statistics → Explore**. Het nagaan of de varianties gelijk zijn zullen we doen aan de hand van een boxplot. Ook deze figuur kunnen we opvragen via **Analyze → Descriptive Statistics → Explore**.

De onafhankelijkheidsassumptie is opnieuw lastig te na te gaan. Indien er geen voor de hand liggende reden is waarom deze assumptie niet voldaan zou zijn, zullen we opnieuw er standaard van uitgaan dat de onafhankelijkheidsassumptie voldaan is.

Voorbeeld ANOVA

Beschouw de volgende onderzoeksvraag:

Is de gemiddelde algemene score verschillend tussen de verschillende richtingen? Indien ja, waar zit dit verschil?

We willen meerdere groepen tegelijk vergelijken, m.a.w. kunnen we gebruik maken van een ANOVA analyse. Om een ANOVA analyse in SPSS uit te kunnen voeren, zullen we eerst de variabele richting omvormen naar een numerieke variabele. Dit lijkt wat vreemd, maar dit is nodig om sommige functies van SPSS te kunnen toepassen. Ga daarvoor naar **Transform → Automatic Recode.** Dubbelklik op richting en geef *richting.num* op als *New Name*. Klik op *Add New Name* en op *OK*.

Eerst moeten we nu nagaan of de verschillende groepen normaal verdeeld zijn en of de varianties gelijk zijn. Ga naar **Analyze → Descriptive Statistics → Explore**, vul bij *Dependent List* alg.score in en bij *Factor List* richting.num. Kies bij de knop *Plots* de optie *Normality plots with tests* en vink *Factor Levels Together* aan bij Boxplots. De andere opties vink je uit of je selecteert *None*.

We vinden voor de richtingen EW, HIR en TEW respectievelijk de p-waarden 0, 292, 0, 627 en 0, 928. De normaliteitsassumptie wordt dus bij geen van de 3 groepen verworpen, wat wordt bevestigd door de QQ-plots. De normaliteitsassumptie lijkt voldaan. Op de boxplot zien we dat de spreiding van de groepen lijkt te verschillen. Deze is veel groter bij de TEW groep dan bij de EW groep. De voorwaarde van gelijke varianties is dus mogelijks niet voldaan. Wegens gebrek aan alternatief, zullen we toch de ANOVA test uitvoeren.

Ga nu naar **Analyze → Compare Means → One-Way ANOVA**. Bij *Dependent List* geef je de Algemene Score in en bij *Factor* kies je nu richting.num. Druk nu op de knop *Post Hoc*, vink **Tukey** aan en druk op *Continue*. De ANOVA test toont een p-waarde van 0, 732, de nulhypothese (de gemiddelden van de groepen zijn gelijk) wordt dus niet verworpen. We besluiten dat er op het 5%-significantieniveau onvoldoende bewijs is om aan te tonen dat de groepen verschillen qua gemiddelde.

Gezien de ANOVA test niet significant is, heeft het geen zin een post hoc analyse uit te voeren. We zullen hier puur ter illustratie deze keer toch naar kijken, we vinden de gecorrigeerde p-waarden 0, 778 (EW vs. HIR), 0, 989 (EW vs. TEW) en 0, 801 (HIR vs. TEW). Dit beaamt de conclusie van de ANOVA test.

3. Multi lineaire regressie

Overzichtelijk in het opgave document.